

WPS 2224

POLICY RESEARCH WORKING PAPER

2224

Growth Forecasts Using Time Series and Growth Models

*Aart Kraay**George Monokroussos*

It is difficult to choose the "best" model for forecasting real per capita GDP for a particular country or group of countries. This study suggests potential gains from combining time series and growth-regression-based approaches to forecasting.



Summary findings

Kraay and Monokroussos consider two alternative methods of forecasting real per capita GDP at various horizons:

- Univariate time series models estimated country by country.

- Cross-country growth regressions.

They evaluate the out-of-sample forecasting performance of both approaches for a large sample of industrial and developing countries.

They find only modest differences between the two approaches. In almost all cases, differences in median

(across countries) forecast performance are small relative to the large discrepancies between forecasts and actual outcomes.

Interestingly, the performance of both models is similar to that of forecasts generated by the World Bank's Unified Survey.

The results do not provide a compelling case for one approach over another, but they do indicate that there are potential gains from combining time series and growth-regression-based forecasting approaches.

This paper — a product of Macroeconomics and Growth, Development Research Group — is part of a larger effort in the group to improve the understanding of economic growth. Copies of the paper are available free from the World Bank, 1818 H Street, NW, Washington, DC 20433. Please contact Rina Bonfield, room MC3-354, telephone 202-473-1248, fax 202-522-3518, email address abonfield@worldbank.org. Policy Research Working Papers are also posted on the Web at <http://www.worldbank.org/research/workingpapers>. The authors may be contacted at akraay@worldbank.org or gmonokroussos@worldbank.org. November 1999. (32 pages)

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the view of the World Bank, its Executive Directors, or the countries they represent.

Growth Forecasts Using Time Series and Growth Models

Aart Kraay

George Monokroussos

The World Bank

1. Introduction

In developed countries, a vast range of forecasting tools have been used to predict growth and other economic variables of interest. In contrast, growth projections for many developing countries are typically based on much more informal techniques. For example, both the World Bank and the International Monetary Fund rely largely on the informed judgement of their country economists to produce forecasts for internal and external use.¹ In this paper, we consider two simple formal models for forecasting growth in a large sample of developed and developing countries: univariate time series models estimated country-by-country, and cross-country growth regressions. The time series models constitute a useful benchmark which illustrates how well forecasts based on extremely limited information (only the history of per capita GDP itself) can perform. The growth regressions are of interest given the vast empirical literature which argues that a significant fraction of the cross-country and time series variation in longer-term growth rates can be explained by a fairly parsimonious set of explanatory variables. A natural question to ask is whether this popular empirical framework has any value for predicting future growth.

We consider the relative forecast performance of two straightforward models. Our time series model is very simple, and models (the logarithm of) real per capita GDP as following a first-order autoregressive process around a broken trend. We estimate this model country-by-country for 112 countries, for two time periods: 1960-1980, and 1960-1990. We then generate out-of-sample forecasts for the remaining years through 1997 based on these two information sets, and compare these forecasts with actual outcomes. Our growth model follows the vast empirical literature spawned by the neoclassical growth model. We estimate a dynamic panel regression of (the logarithm of) real per capita GDP on itself lagged five years, and a number of lagged explanatory variables which proxy for the steady-state of the neoclassical growth model and capture the effects of various policies on long-run growth: investment, population growth, trade openness, inflation, and the black market premium. We estimate this model using non-

¹ The World Bank's Unified Survey projections, and the IMF's World Economic Outlook projections are produced in this way. Both organizations also use large macroeconomic models: the World Bank's Global Economic Model (GEM) is used to produce forecasts appearing in the Bank's annual Global Economic Prospects publication, and the IMF maintains MULTIMOD for research and simulation purposes.

overlapping quinquennial averages of data over the same two periods as for the time series model (although for a somewhat smaller sample of countries as dictated by data availability), and then generate forecasts for the remaining years in the sample which can be compared to actual outcomes. In order to benchmark the forecasts generated by these models against current practice, we also make some comparisons with long-term forecasts produced by the World Bank's Unified Survey in 1990. However, our primary interest is in the relative performance of the time series and growth models.²

We assess the out-of-sample forecast performance of these models using standard summary statistics which capture their bias and mean squared error. These statistics suggest small median (across countries) differences in forecast performance of the alternative models, which vary with the forecast horizon. For example, there is some evidence -- consistent with our priors -- that the mean squared error of growth regression based forecasts is smaller at long forecast horizons (five years or more). However, these differences in median forecast performance are typically very small relative to the cross-country dispersion in forecast performance, casting doubt on the significance of observed "typical" differences. The relative performance of the alternative forecasting models is also very unstable over time within countries. We test for and do not reject the null hypothesis that the past relative performance of the growth model and the time series model in a particular country is independent of the future relative performance of the two models in that country.

These results indicate that neither forecasting model dominates, both across countries and within countries over time. Rather than attempt to choose a single "best" forecasting model, we instead ask whether there is value in combining the forecasts of alternative models. We implement forecast encompassing tests and find evidence that these approaches can "learn from each other", in the sense that the forecasts from both models are jointly significant in explaining actual outcomes. This is especially true at shorter horizons, and it suggests that there are potential benefits from combining these forecasts in some way to arrive at a superior overall method.

² For a more systematic assessment of the quality of World Bank forecasts, see Ghosh and Minhas (1993), and Verbeek (1999). Artis (1996) does the same for the IMF's short-term forecasts.

The remainder of this paper proceeds as follows. In the next section, we present the two models used to produce growth forecasts, and note the similarities and differences between them. In Section 3, we examine the cross-country performance of these forecasts using various summary statistics. In Section 4, we illustrate the results of our forecast encompassing tests, and consider whether a combined forecast can outperform either of the two alternatives. We also briefly consider whether the absolute performance of either model is adequate. Section 5 offers some concluding remarks.

2. Forecasting Models

In this section we describe the simple time series and growth models we use to forecast real per capita GDP in a large sample of developed and developing countries.

2.1. Time Series Forecasts

For each country, we estimate a very simple first-order autoregressive process around a linear trend, allowing for the possibility that the trend of the series changes once within the estimation period. In particular, we assume that the logarithm of real per capita GDP in country i at time t , y_{it} , is described by the following process:

$$(1) \quad y_{it} = \rho_i \cdot y_{i,t-1} + \delta_{it} + \varepsilon_{it}$$

The trend term δ_{it} is a linear function of time, and both the slope and the intercept term may change at a date T within the estimation period, i.e. $\delta_{it} = \mu_i + \theta \cdot D_T^\mu + \beta \cdot t + \gamma \cdot D_T^\beta$, where D_T^μ is a dummy variable taking on the value 1 if $t > T$ and zero otherwise, and D_T^β is a dummy variable taking on the value $t - T$ if $t > T$ and zero otherwise. The two dummy variables pick up a shift in the deterministic component of output that occurs in year T . The date of the trend break, T , is determined endogeneously, using the procedure of sequential Wald tests suggested by Vogelsang (1997).³ At the estimation stage, we do not need to make strong assumptions about the properties of the error term. However, for the purposes of formal tests of model performance, it will be useful to assume that the error term is independent over time and is normally distributed with variance σ_i^2 .

In order to evaluate the forecasting performance of this model, we divide the sample period in two at a particular year t . We then estimate Equation (1) using the data available until this year t , and then use the model to forecast the log-level of per capita

³ However, we do not pre-test for a trend break, i.e. we allow for a trend break at time T even if this break is not statistically significant. There is some evidence that forecasts based on pre-tested models perform better than either of the alternative models that are being pre-tested (Diebold and Kilian (1999) perform Monte Carlo experiments, and Stock and Watson (1998) show this empirically in a large-scale comparison of many forecasting models of various macroeconomic aggregates for the United States). This suggests that the forecasting performance of both the time series model and the growth model might be improved by pretesting.

GDP for each subsequent year. In particular, if we divide the sample in two at year t , our forecast of per capita GDP for each subsequent year is:

$$(2) \quad \hat{y}_{i,t+s|t} = \hat{\rho}_i^s \cdot y_{it} + \hat{\delta}_{i,t+s}$$

where $\hat{y}_{i,t+s|t}$ denotes the forecast of $y_{i,t+s}$ based on information available at time t and $\hat{\rho}_i$ and $\hat{\delta}_{i,t+s}$ are the parameter estimates for country i based on its data available through year t . Ignoring the uncertainty associated with the parameter estimates, i.e. assuming the parameters of the model are known, the corresponding forecast error is:

$$(3) \quad e_{i,t+s|t} = \sum_{h=0}^{s-1} \rho_i^s \cdot \varepsilon_{i,t+s-h}$$

The variance of this error term can be used to construct the ex ante forecast confidence intervals associated with each forecast, which will depend on the autoregressive parameter, ρ_i , and the variance of the error term, σ_i^2 . Replacing these with their estimates yields the usual ex ante forecast confidence intervals.⁴

Our data consists of a panel of 112 countries for which a complete time series on real per capita GDP adjusted for differences in purchasing power parity is available over the period 1960-1997.⁵ We estimate this model twice for each country, once using data over the period 1960-1980, and once over the period 1960-1990. We then generate forecasts of real per capita GDP for the remaining years through 1997 for each country, and compare these forecasts with the actual realizations of per capita GDP for each country.

⁴ In particular, a 90% forecast confidence interval extends $\pm 1.64 \cdot \sqrt{\sum_{h=0}^{s-1} \hat{\rho}_i^{2s} \cdot \hat{\sigma}_i^2}$ around the forecast itself.

2.1. Forecasts based on cross-country growth regressions

The cross-country growth regressions we consider differ from the simple time-series model in three important respects. First, unlike the time series model, the growth regression has a clear theoretical motivation which permits the inclusion of country-specific explanatory variables into the model. Second, the growth regression is typically estimated using longer averages of data over non-overlapping periods rather than annual observations. Third, many of the parameters of the growth model are restricted to be equal across countries. We discuss each of these differences in turn.

The theoretical motivation for many cross-country growth regressions is the prediction of the neoclassical growth model for the dynamics of per capita output around its steady state. A fundamental prediction of this model is that per capita GDP growth declines as per capita GDP approaches its steady-state level, i.e.

$$(4) \quad y_{i,t} - y_{i,t-1} = (1 - \rho_i) \cdot (y_{it}^* - y_{it})$$

where y_{it}^* denotes the steady state of country i at time t (note that the steady state may itself evolve over time), and ρ_i denotes the annual rate of convergence in country i . Adding an error term which captures deviations between this model of the long run and reality, and rearranging, yields an empirical specification which is very similar to the time series model in Equation (1):

$$(5) \quad y_{it} = \rho_i \cdot y_{i,t-1} + (1 - \rho_i) \cdot y_{it}^* + \varepsilon_{it}$$

This illustrates the first difference between the time series model and the growth model. In the growth model, growth theory provides variables that can serve as proxies for the steady state, y_{it}^* , and hence permit empirical estimation of Equation (5). In contrast, the time series model can be thought of as proxying the steady state log-level of income for each country with a country-specific trend (with a possible break).

⁵ The data is drawn from the Penn World Table Version 5.6 (RGDPCH) and is extended through 1997 using World Bank constant price local currency growth rates.

The second difference between the two models is that the growth regression is typically estimated using (possibly a panel) of long-run averages of both GDP and the proxies for the steady state. To see the consequences of this, we can iterate Equation (5) forward for T periods, corresponding to a growth regression estimated using T -year average growth rates:

$$(6) \quad y_{i,t+T} = \rho_i^T \cdot y_{i,t} + \sum_{h=0}^{T-1} \rho_i^h \cdot ((1 - \rho_i) \cdot y_{i,t+T-h}^* + \varepsilon_{i,t+T-h})$$

To empirically implement this equation, we require proxies for the (possibly changing) steady state of the economy between periods t and $t+T$. These are usually taken to be averages over the same period of variables such as population growth, the investment rate, various measures of policies which affect the long-term growth prospects of a country, and possibly an unobserved country-specific effect. In particular, it is typically assumed that $y_{i,t+T-h}^* = \mu_i + \beta_i' x_{i,t}$, where $x_{i,t}$ is a vector of such proxies for the steady state and μ_i is an unobserved country-specific effect. Inserting this into Equation (6) gives the standard cross-country growth regression:

$$(7) \quad y_{i,t+T} = \rho_i^T \cdot y_{i,t} + (1 - \rho_i^T) \cdot (\mu_i + \beta_i' x_{i,t}) + v_{i,t}$$

where $v_{i,t} = \sum_{h=0}^{T-1} \rho_i^h \cdot \varepsilon_{i,t+T-h}$ is a composite error term reflecting all of the annual shocks that occurred between t and $t+T$.

The third difference between the growth model and the time series model is that the growth model is estimated pooling data for many countries and restricting most of the parameters in Equation (7) to be the same across countries, while the time series model is estimated country-by-country and imposes no such restrictions. In particular, we estimate the growth model in Equation (7) using a panel of non-overlapping quinquennial averages, restricting ρ and β to be the same across countries. We treat the country-specific effects μ_i as unobserved, and estimate the model using the GMM system estimator for dynamic panels suggested by Arellano and Bover (1995). This

method is superior to simple pooled OLS or IV estimation of Equation (7) because it allows for a consistent treatment and estimation of the individual effects.⁶

As with the time series model, we estimate the growth regression in Equation (7) twice, using data available through 1980, and data available through 1990, and then project real per capita GDP forward for the remaining years in the sample using the estimated parameters as follows:

$$(8) \quad \hat{y}_{i,t+s|t} = \hat{\rho}^s \cdot y_{it} + (1 - \hat{\rho}^s) \cdot (\hat{\mu}_i + \hat{\beta}' x_{it})$$

Again ignoring the uncertainty associated with the parameter estimates, this results in exactly the same forecast error as for the time series model, except that the autoregressive parameter ρ is now the same for all countries:

$$(9) \quad e_{i,t+s|t} = \sum_{h=0}^{s-1} \rho^h \cdot \varepsilon_{i,t+s-h}$$

This expression can be used to construct ex ante forecast confidence intervals in the same way as for the time series model.

We implement the growth model using an unbalanced panel of non-overlapping quinquennial data over the period 1961-1995, $T=5$. The vector of explanatory variables x_{it} consists of a constant, the logarithm of the investment rate, the logarithm of the population growth rate, the logarithm of one plus the CPI inflation rate, the logarithm of one plus the black market premium, and the share of trade in GDP. The first two variables follow the predictions of the textbook Solow model. The last three variables can be interpreted as summary indicators of policy. We begin with the same sample of countries as with the time series models, but we can only estimate the growth regressions for a somewhat smaller sample due to missing values for some of the explanatory variables.

⁶ We also considered the forecasting performance of a growth model estimated using OLS, which has the convenience of much simpler implementation. Despite the theoretical advantages of the dynamic panel model, the ex post forecast performance of the dynamic panel model is not consistently better than that of the simple OLS model.

We estimate (7) using averages of the variables in x_{it} in the five years prior to t . We do this because when we turn to the growth forecasts in Equation (8), we can generate forecasts without also having to forecast each of the explanatory variables in the growth regression. At the estimation stage, this approach also has the advantage of alleviating some of the concerns about the endogeneity of contemporaneous values of the “growth determinants” in most empirical growth specifications. The disadvantage of this is that this growth regression does not fit the data as well as a regression which uses contemporaneous values of the explanatory variables: the average of a growth determinant over $(t, t+T)$ is typically a better explainer of growth over $(t, t+T)$ than is the average of the same variable over $(t-T, t)$. However, forecasts of real per capita GDP based on such a model would also require forecasts of each of the explanatory variables in the growth regression.⁷

The results of estimating Equation (7) for the two information sets are shown in Table 1. As a benchmark, we report estimates using OLS on the pooled sample of five-year averages, and also our preferred specification based on the system GMM estimator for dynamic panel data. The results are broadly consistent with both intuition and existing results. The lagged level of income enters significantly with a coefficient less than one in all cases, and is smaller (implying a higher estimated rate of convergence) for the GMM estimator. Population growth and investment are always highly significant, and the magnitude of the estimated coefficients are reasonably stable. Openness and the black market premium generally enter with the expected signs, but are not consistently significant. Unfortunately inflation often enters with a perverse positive sign, although it is only significant when it is negative. The less-than-stellar performance of the policy variables in the growth regression is somewhat disappointing, and is in part due to the fact that these are lagged policy variables, rather than contemporaneous.

In summary, the time series model and the growth model can be thought of as special cases of the same general model in which the log-level of real per capita GDP

⁷ As a robustness check, we also estimated the growth model using contemporaneous values of the explanatory variables, and then generated forecasts by inserting the actual future values of the explanatory variables into the forecasting equation. This corresponds to the unrealistic assumption that the forecaster has perfect foresight for all of the explanatory variables when producing growth forecasts. Not surprisingly, (a) the growth model fits somewhat better in sample, and (b) the forecasts generated by this model perform somewhat better, although not by much.

follows a first-order autoregressive process around a trend. In the time series model, the trend is modelled as a simple function of time with at most one shift. In the growth model, the trend term is interpreted as the steady state of the neoclassical growth model, and is proxied by variables suggested by the theory. As a result, the forecasts generated by the growth model are based on more information than the time series model, since they incorporate proxies for the steady state for each country. Although in general one would expect that this should lead to superior forecasts, this advantage is to some extent offset by the fact that the growth model forces the parameters of the model to be the same across countries, while the time series model allows them to differ across countries⁸. Since the balance of these two effects is ambiguous, there is no a priori reason to prefer one method over the other.

⁸ Attempting separate within country growth regressions would probably be of limited usefulness because of insufficient within-country variation in determinants of long-run growth over our sample period.

3. Results

In this section, we provide a description of the forecasting performance of the time series model and growth model. We begin by looking at the how the various forecasts fare for a few specific countries, and then provide a number of descriptive statistics which summarize the ex post performance of these models for a large number of developed and developing countries. Finally, we provide some comparisons between both these models and those reported in the World Bank's Unified Survey.

3.1. A Look At Individual Country Forecasts

It is interesting to begin by looking at forecasts for a few selected countries, in order to get a sense of why different methodologies lead to different forecasts. The top left corner of Figure 1 shows the actual log-level of per capita GDP for Nigeria, as well as forecasts for the period 1981-1996 based on both models. In the case of Nigeria, the growth model clearly outperforms the time series model. The time series model identifies a trend break in per capita GDP around 1970 for Nigeria, and then extrapolates the trend growth during the 1970s into the 1980s and 1990s. As a result, it misses entirely the five years of negative growth during the first half of the 1980s and subsequent stagnation that actually occurred. In contrast, the growth regression fares much better as it in part accounts for Nigeria's worse policy and structural determinants in the second half of the 1970s which had predictive power for Nigeria's subsequent performance.

However, it would be misleading to conclude from Figure 1 that policy-based growth regressions are in general much better at forecasting growth. In bottom left panel of Figure 1, we plot the opposite case of the Netherlands. Here, the growth model performs worse than the time series model, predicting significantly lower growth during the 1980s and 1990s than actually occurred. There are also many cases where neither model does very well. For example, the top right panel of Figure 1 plots the same graph for Argentina, but this time using forecasts based on information available in 1990. Here the growth model and the time series model do equally poorly in predicting the turnaround in Argentina during the 1990s relative to the 1980s. Finally, there are

countries such as the United States where both models perform more or less equally well (see the bottom right panel of Figure 1).

It is also useful to distinguish between the two models in terms of their ex ante forecast confidence intervals. To avoid cluttering the graphs excessively, we show these intervals for Argentina only, in Figure 2. The most striking feature of Figure 2 is that these forecast confidence intervals are much larger for the growth model than for the time series model. In particular, the 90% forecast interval for the growth regression after five years is around ± 0.2 , which translates into a 90% confidence interval for the average annual growth forecasts over this period of around $\pm 3.7\%$ per year ($(1.20)^{(1/5)} - 1 = 0.037$). In contrast, for the time series model the 90% forecast interval is around ± 0.03 , which translates into a 90% confidence interval for the average annual growth forecasts over this period of around $\pm 0.6\%$ per year ($(1.03)^{(1/5)} - 1 = 0.0059$). This difference in the ex ante confidence associated with the forecasts reflects the fact that the in-sample fit of the time series model is much better than the in-sample fit of the growth model.⁹

The main lesson from this first look at the data is that it is difficult to say a priori which forecasting method will do best. We explore these issues more systematically below by looking at the summary statistics of forecast quality for all of the countries and all of the forecasts in our sample.

3.2 Cross-Country Comparisons of Forecast Models

We now turn to a more formal and systematic evaluation of the ex post performance of the forecasts generated by these methods. We use two simple statistics which capture the bias and mean squared error of the forecasts.¹⁰ For each country i , we measure the bias of an h -period ahead forecast as the cumulative sum of the

⁹ Since the growth model forces the autoregressive parameter ρ and the variance of the error term σ to be the same across all countries, the forecast confidence intervals are the same for all countries (recall Equation (9)). In contrast, the forecast confidence intervals for the time series model vary across countries, as these estimated parameters also vary across countries. Neither set of confidence intervals reflects the uncertainty associated with the estimates of the parameters themselves.

¹⁰ We do not use “rationality” tests to evaluate forecasts, as is often done in the literature. In this literature, a forecast is accepted as “rational” if there are no variable available at the time that the forecast is made which have explanatory power for the subsequent forecast errors. In practice, these tests are of limited usefulness in selecting between alternative forecasting models since there is a potentially unlimited number of explanatory variables which need to be considered before a forecast can be shown as rational.

forecast errors. In order to make this comparable across countries, we scale this sum by the actual outcomes, resulting in the following cumulative forecast error statistic:

$$(10) \quad CFE_{i,h|t} = \frac{\sum_{s=1}^h (\hat{y}_{i,t+s|t} - y_{i,t+s})}{\sum_{s=1}^h y_{i,t+s}}$$

All other things equal, it is natural to prefer forecasts with cumulative forecast errors near zero.

Similarly, for each country i we measure the variability or precision of an h -step ahead forecast using the sum of squared forecast errors. To make this comparable across countries, we scale it by the sum of squared actual outcomes, resulting in what is known as the Theil U-statistic:

$$(11) \quad TU_{i,h|t} = \frac{\sum_{s=1}^h (\hat{y}_{i,t+s|t} - y_{i,t+s})^2}{\sum_{s=1}^h y_{i,t+s}^2}$$

All other things equal, we would prefer forecasting methods with low Theil U-statistics, since the variability of the forecast errors is low relative to the variability of real per capita GDP.¹¹

For each country, we calculate the CFE and TU for both forecasting models, based on information available through 1980, and through 1990, for every possible forecast horizon. In Figures 3-6 we provide a graphical overview of these many summary statistics of forecast performance. In Figure 3 we consider the sample of 73 countries for which we are able to produce forecasts using all five methods in 1980.¹² We plot time on the horizontal axis, and on the vertical axis, we plot the median across countries of the two measures of forecast quality discussed above, the CFE (upper

¹¹ As is well known, the MSE of a forecast can be written as the sum of the variance of the forecast errors plus the bias squared. As such it reflects a particular weighting of bias and precision in assessing forecast quality. However, for many purposes the bias in a forecast is of independent interest. For this reason we report both the cumulative forecast error and the Theil U statistic for each country.

panel), and the TU (lower panel). We report the medians rather than the means, since for some countries, one model or the other can deliver “crazy” forecasts resulting in very large TUs or CFEs (in absolute value).

In the upper panel of Figure 3, the time series model and the growth model have very similar performance in terms of the CFE statistic, which measures the bias in forecasts. Both models significantly over-predict real per capita GDP -- and do so increasingly over time¹³. This occurs because, on average, both models do a rather poor job of predicting the worldwide slowdown in growth during the first half of the 1980s. To interpret the magnitude of this bias, recall that the vertical axis measures logarithm of real per capita GDP. Since, for example, the cumulative median bias in the level of forecasted real per capita GDP after five years is around 7% of per capita GDP, this translates into an upward bias in average annual growth forecasts over this period of around 1.4% per year ($(1.07)^{(1/5)} - 1 = 0.0136$).

Turning to the TU statistics in the lower panel, there is a somewhat clearer distinction between the two models. At all forecast horizons, the growth model delivers a lower variability of forecast errors, as reflected in lower TU statistics. This gap between the two models widens over time, suggesting that the relative performance of the growth model is better for longer-term growth forecasts. In contrast, at short horizons, e.g. less than 5 years, the performance of the two models is rather similar. Somewhat surprisingly, the relative performance of these models according to both criteria is similar in a smaller sample of 59 developing countries for which we have forecasts from both models. The graphs summarizing these results are omitted for brevity.

In Figure 4 we do the same exercise, but for forecasts based on information available through 1990, using a slightly larger sample of 82 countries. As in the 1980s, the forecasts of both models are on average biased upwards, although less so than in the 1980s forecasts (note that the units of the vertical axis are very different in Figures 4 and 3). The median bias in the forecasts is never greater than 0.4% of GDP. In

¹² Although we can produce the time series forecasts for all 112 countries in our data set, we have complete data on all of the explanatory variables required for the growth regression in 1980 for only 73 countries, and for 82 countries in 1990.

contrast to the 1980s forecasts, the growth model does somewhat better than the time series model, both in terms of bias (see the CFEs in the upper panel) and in terms of variability (see the TUs in the lower panel).

In Figure 5, we repeat the information in Figure 4, but for a smaller set of only developing countries for which we also have the Unified Survey forecasts produced by the World Bank in 1990.¹⁴ During this period, the performance of the Unified Survey forecasts was remarkably similar to that of the other two models, both in terms of bias and mean squared error. Interestingly, there is little evidence that the Unified Survey's long term forecasts are biased upwards during this period. This is in contrast to other findings that World Bank forecasts are typically over-optimistic (Ghosh and Minhas (1993)). However, given the large differences in the performances of forecasts based on different information sets, it is premature to conclude that this finding is general.

Thus far, we have seen that the median (across countries) performance of all three models considered here are quite similar, with the growth regression perhaps having a slight advantage over the other two alternatives. A natural question is whether any of the differences in median performance of these models are either economically or statistically significant. One way to answer this question is to look at the entire cross-section of CFE and TU statistics at every forecast horizon, in order to obtain a sense of whether the differences in medians are representative. We do this in Figure 6, for the CFE statistics in Figure 5. In the first panel, we reproduce the first panel of Figure 5, but add vertical bars to the CFEs for the time series model indicating the interquartile range of the cross-sectional distribution of the CFE statistics. In order not to clutter the graph excessively, in the next two panels we report the same information, but instead for the CFEs of the growth model, and the Unified Survey forecasts, separately. The most striking feature of these graphs is that the cross-sectional distribution of these statistics is extremely dispersed. For each model, the interquartile range of the CFE statistics swamps any differences in the medians of these statistics, suggesting that differences in

¹³ A potentially useful thing to do (in order to decrease this observed bias) would be a "Dynamic Estimation" of both of our models, where estimation is achieved by minimizing the in-sample counterpart of the desired multi-step ahead forecast horizon, which thus produces a different parameter estimate per forecast horizon. (And thus avoids raising our estimated parameters to powers (see equations (2), (8)), which could seriously exacerbate bias problems for large forecast horizons.)

median performance are highly unlikely to be of statistical or practical relevance. Similar graphs indicating the cross-country dispersion in the TU statistics (not shown for brevity) lead to a similar conclusion that the cross-country variation in model performance is large relative to the differences in median performance.

Finally, we ask whether the relative performance of the various models is stable over time for a particular country. This question is relevant if one is interested in producing forecasts for a particular country and it is necessary to choose one model over another. In this case, it would be useful to know whether the fact that, for example, the growth model outperformed the time series model for that country in the past is a good predictor of the future relative performance of the two models. To answer this question, we focus on the five-year ahead forecasts of real per capita GDP generated in 1980 and in 1990, for both the time series and the growth model. We then use the Theil-U statistics for each country to assess the relative performance of the two models in each of the two forecasting periods. We summarize the results with a two-way classification of countries, indentifying which model dominated the other (in the sense of having a lower TU statistic) in each of the two years. The results of this calculation are summarized in Table 2 for the set of 73 countries common to both samples. Unconditionally, the probability that the time series model outperforms the growth model is around 0.44, corresponding to 33 out of 73 countries for the period 1981-85, and 31 out of 73 countries in the period 1991-95.

The main question of interest is whether the time series model consistently outperforms the growth model in the same countries over time. The remainder of the table indicates that this is not the case. In only 13 out of 73 countries does the time series model outperform the growth model in both periods, and the converse occurs in only 22 out of 73 countries. For the remaining countries, one model fares relatively well in the one period but not so in the other. In fact, a chi-squared test of the null hypothesis that relative performance over the period 1981-85 is uncorrelated with relative performance over the period 1991-95 (i.e. of the independence of the rows and columns) yields a p-value of 0.63. This comfortably rejects the notion that, within countries, past relative forecast performance is any guarantee of future forecast performance.

¹⁴ These forecasts are taken from the 1991 Unified Survey, so that most data through 1990 would have been available at the time these forecasts were made. In this version of the Unified Survey, 10-year

4. Implications

In this section we take up two questions suggested by the results of the previous section. First, given that both the time series and the growth models perform comparably, is it possible to combine them in some way to arrive at better forecasts? Second, does either model perform well in absolute (as opposed to relative) terms?

4.1 Can Alternative Forecasting Models “Learn” From Each Other?

Thus far, we have seen that there is little clear evidence suggesting that we should select one model over another as a forecasting tool. Rather than restrict ourselves to selecting one model over another, a more constructive approach is to ask whether some combination of models leads to better forecasts. We do this using tests of forecast combination and encompassing. Intuitively, these tests ask whether alternative forecasting models can “learn” from each other. If they do, this suggests that superior forecasts can be obtained by combining the two models in some way.

Formally, suppose that both the time series and the growth model generate forecasts that are informative for future real per capita GDP, so that we can write future real per capita GDP as a linear combination of the two forecasts plus an error term:

$$(12) \quad y_{i,t+s} = \alpha + \beta \cdot \hat{y}_{i,t+s|t}^{GR} + (1-\beta) \cdot \hat{y}_{i,t+s|t}^{TS} + u_{i,t+s}$$

where the superscripts TS and GR differentiate between the forecasts of the time series and growth models.¹⁵ We can then test the null hypothesis that the time series model forecast-encompasses the growth model by testing the null that $\beta=0$. The intuition for this test is straightforward, since it simply asks whether the variation in the growth-regression based forecasts that is orthogonal to the time series forecasts has any useful explanatory power for actual outcomes. If it does not, then the time series model “encompasses” the growth model in the sense that the growth model’s forecasts provide

average annual growth projections were produced, and we use these average annual growth rates to forecast growth for every year through 1996.

no additional predictive power for actual output. Conversely, we can test whether the growth model encompasses the time series model by testing whether $\beta=1$.

We implement these tests for the time series and growth models discussed above as follows. For every forecast horizon, we estimate Equation (12) cross-sectionally, and test the two null hypotheses that each model encompasses the other.¹⁶ The results of the cross-sectional regressions are shown in Table 3. In general, we reject the null hypothesis that either model encompasses the other at short forecast horizons, suggesting that both models can benefit from incorporating features of the other. Interestingly, as the forecast horizon increases, we do not reject the null hypothesis that the growth model encompasses the time series model (i.e. $\beta=1$), but not the converse. This is consistent with the notion that long-run growth regressions do a better job of predicting long-term growth.

These results strongly suggest that there are benefits to combining the information from both forecasting models. However, it is less clear exactly how this should be done. A simple approach would be to use some weighted average of forecasts from the two models. There is some empirical evidence in other contexts that weighted (or even unweighted) averages can outperform the components of the average (e.g. Stock & Watson (1998)). However, Diebold (1989) stresses that there is no guarantee that this will be the case in general. Since neither model individually does a very good job at capturing the “true” underlying data generating process, there is no reason to believe that a combination of the two will do so on a consistent basis. For example, the large negative intercepts in the encompassing regressions for the 1980s forecasts reflect the large ex post positive bias in these forecasts. However, if we were to use this information to systematically lower all growth forecasts for the 1990s (as the encompassing regression might suggest), we would have ended up significantly underpredicting growth in the 1990s.

¹⁵ The restriction that the coefficients on the two models sum to one is not essential. Estimating these encompassing regressions without such a restriction leads to very similar results.

¹⁶ We also carry out these tests using the time series of forecast errors for each country, from the forecasts based on information available through 1980. For each country, we estimated rolling regressions, starting with the time series of the first 10 forecast errors over the period 1981-90, and continuing through the entire time series of errors through 1997. The results of this exercise were consistent with the cross-sectional results. At short horizons, neither model encompassed the other. However, it was more likely that the growth model encompassed the time series model at long horizons than the other way around.

Instead, a more compelling approach is to combine the information sets on which the forecasts are based, rather than combine the forecasts themselves (Clements and Hendry (1998), Diebold (1989)). For example, a natural way to combine the two models would be to consider a hybrid time series model which (a) includes additional explanatory variables that our time series model omits, and (b) relaxes the restriction of the growth models that the parameter estimates are equal across countries. An example of such a modelling strategy might be a non-structural vector autoregression in several key macroeconomic variables, estimated country-by-country. Forecasts based on such a combined information set are more likely to encompass alternatives since they making optimal use of all of the (useful) available information in both information sets.

4.2 Formal Tests of Predictive Failure

Thus far, our emphasis has been on comparing the relative performance of alternative forecasting models. We now turn to a rather different question: is the absolute performance of these forecasting models adequate? Alternatively, is the ex post performance of these forecasting models good enough (or bad enough) that we should continue to use some combination of these models (or search for other forecasting models)?

In principle, this question can be answered using the test of predictive failure developed by Box and Tiao (1976). Intuitively, this test asks whether the deviations between forecasts and actuality are large relative to the forecast confidence intervals generated ex ante. To see how this works, consider the case of Argentina, where we have already seen the forecast confidence intervals in Figure 2. For the case of a one-step ahead forecast, the Box-Tiao test simply asks whether the actual outcome of real per capita GDP falls within the ex ante confidence interval generated by the forecaster. If it does, we do not reject the null hypothesis that the forecasting model is correctly specified, since, roughly speaking, the actual outcome fell within the range that was expected a priori. This, however, should not be taken as an endorsement of the model either, since failure to reject the null hypothesis may simply reflect large ex ante

confidence intervals generated by a model that fits very poorly within-sample. If in contrast the actual outcome falls outside the range predicted by the forecaster, the Box-Tiao test rejects the model.

In the case of Argentina, we see that for the one-year ahead forecast, i.e. the forecast of real per capita GDP in 1991 based on information in 1990, falls inside the confidence interval of the growth model, but outside the confidence interval of the time series model. The Box-Tiao test therefore suggests that we should reject the time series model, but not the growth model, as a forecasting tool. We have implemented the version of the Box-Tiao test appropriate for multi-step forecasts for all countries, and we find a similar pattern to that observed in Argentina.¹⁷ For the great majority of countries, the Box-Tiao test suggests that we should reject the time series model. For the growth model, the Box-Tiao test rejects the growth model for about half of the countries, and fails to reject for the other half.¹⁸

Is this conclusion warranted? As noted above, the Box - Tiao test may fail to reject a model simply because the model is very imprecise ex ante. Indeed, in our context, and as is clear from Figure 2, the main reason why the growth model is not rejected by the Box-Tiao test is because the ex-ante forecast confidence intervals associated with this model are so large as to render the accompanying growth forecasts virtually meaningless. We have already noted that the growth regression typically generates ex ante forecast intervals of plus or minus four percent per year around a typical 5-year ahead growth forecast. This spans almost the entire range of actual growth performance in most periods. Conversely, the time series model fares relatively poorly according to the Box-Tiao criterion for assessing the significance of predictive failure because the ex ante confidence intervals associated with the time series model are far smaller than those of the growth model. This reflects the fact that the time series model tends to "over-fit" the data in sample. Given that the parameters of the time series model are typically rather imprecisely estimated (especially the date of the trend break), forecast confidence intervals which take this into account would give a more

¹⁷ In particular, Box and Tiao (1976) show that for a process with Gaussian innovations, $e_t' \Omega^{-1} e_t$ has a $\chi^2(h)$ distribution, where e_t is the $h \times 1$ vector of forecast errors through h and $\Omega = E[e_t e_t']$. Given the expressions for the forecast errors in Equations (3) and (9) and the corresponding estimates of ρ_t , it is possible to obtain an estimate of Ω and compute the appropriate test statistic.

reasonable picture of the ex ante uncertainty of this model's forecasts. As a result, the Box-Tiao test would be less likely to reject this model as a forecasting tool.

¹⁸ Recall that the forecast confidence intervals are the same across countries for the growth model. For the time series model, there is some variation, but since in general the time series model fits the data very well in-sample, the forecast confidence intervals tend to be quite small for all countries.

5. Conclusions

In this paper, we have considered the relative performance of two simple forecasting models for real per capita GDP in a large sample of developed and developing countries: a univariate time series model for real per capita GDP, and a cross-country growth regression model. The most striking finding of this paper is that neither model clearly dominates as a forecasting tool. Median (across countries) differences in the forecasting performance of the two models are typically very small relative to the cross-country variation in relative model performance. Moreover, both absolute and relative model performance is very unstable over time. Both models significantly overpredict growth in the 1980s, but do not in the 1990s. Within countries, past relative forecast performance is uncorrelated with future relative forecast performance.

These results indicate that it is very difficult to choose the “best” forecasting model for a particular country or group of countries. Instead of attempting such a choice, our results suggest that there are potential benefits from combining the two forecasting methodologies. Forecast encompassing tests indicate that the forecasts of both models are jointly informative for actual outcomes, especially at shorter horizons. A natural way to proceed would be to combine the information sets from the two models in some way. In particular, vector autoregressions in a small set of key macroeconomic variables, estimated country-by-country, may improve over the forecast performance of both models. The advantage of such an approach over the univariate time series models is that it draws on a larger information set. This approach can potentially also improve over forecasts based on cross-country growth regressions by relaxing the restrictive assumption that the parameters of the model are equal across countries.

References

- Arellano, Manuel and Olympia Bover (1995). "Another Look at the Instrumental Variable Estimation of Error-Components Models". *Journal of Econometrics*. 68:29-51.
- Artis, Michael (1996). "How Accurate are the IMF's Short Term Forecasts? Another Examination of the World Economic Outlook". Manuscript: European University Institute.
- Box, G.E.P., and G.C Tiao (1976), "Comparison of Forecast and Actuality", *Applied Statistics*, 25, 195-200.
- Clements, M, and David Hendry (1998), "Forecasting Economic Time Series", CUP.
- Diebold, F (1989), "Forecast combination and encompassing: Reconciling two divergent literatures", *International Journal of Forecasting*, 5, 589-592.
- Diebold, F, and L. Kilian (1999), "Unit root tests are useful for selecting forecasting models", NBER WP 6928.
- Ghosh, Atish and Tanya S. Minhas (1993). "How Good Are World Bank Projections? A Post-Mortem for the Period 1975-1991.
- Stock, J.H., and M. Watson (1998), "A Comparison of Linear and Nonlinear Univariate Models for Forecasting Macroeconomic Time Series", NBER Working Paper 6607.
- Verbeek, Jos (1999). "The World Bank's Unified Survey Projections: How Accurate Are They? An Ex-Post Evaluation of US91-US97". World Bank Policy Research Department Working Paper No. 2071.
- Vogelsang, T.J. (1997), "Wald-Type Tests for Detecting Breaks in the Trend Function of a Dynamic Time Series", *Econometric Theory*, 13, 818-849.

Table 1: Growth Regression ResultsDependent variable is $\ln(\text{real per capita GDP})$

	1980		1990	
	OLS	GMM	OLS	GMM
Constant	--	2.227 (1.168)*	--	-0.526 (0.410)
Lagged $\ln(\text{real per capita GDP})$	0.923 (0.019)***	0.748 (0.118)***	0.940 (0.015)***	0.862 (0.053)***
$\ln(\text{Population Growth} + 0.05)$	-0.253 (0.099)***	-0.247 (0.224)	-0.330 (0.071)***	-0.683 (0.113)***
$\ln(\text{Investment/GDP})$	0.076 (0.021)***	0.324 (0.125)***	0.053 (0.017)***	0.076 (0.044)*
$\ln(\text{Inflation})$	0.069 (0.115)	0.295 (0.230)	-0.166 (0.068)***	0.003 (0.132)
$(\text{Exports} + \text{Imports})/\text{GDP}$	0.06 (0.033)*	-0.281 (0.191)	0.012 (0.020)	0.046 (0.051)
$\ln(1 + \text{Black Market Premium})$	-0.076 (0.064)	-0.124 (0.149)	-0.061 (0.037)*	-0.075 (0.040)*
P-Value for Sargan Test of ODR		0.859		0.238
P-Value for no SOSC		n/a		0.716

Table 2: Persistence of Relative Forecast Performance

		1991-95		
		TS Dominates	GR Dominates	Total
1981	TS Dominates	13	20	33
-1985	GR Dominates	18	22	40
	Total	31	42	73

P-Value for Chi-Squared Test of Independence: 0.63

Notes: This table reports the relative forecast performance of the time series model (TS) and the growth model (GR), for 5-year growth forecasts for 1981-85 and 1991-95. The cells of the table indicate the number of countries for which the TU statistic of the TB model is lower than that of the GR model (TS Dominates), and conversely the number of countries for which the TU statistic of the GR model is lower (GR Dominates) during the indicated forecast periods.

Table 3: Forecast Encompassing Tests

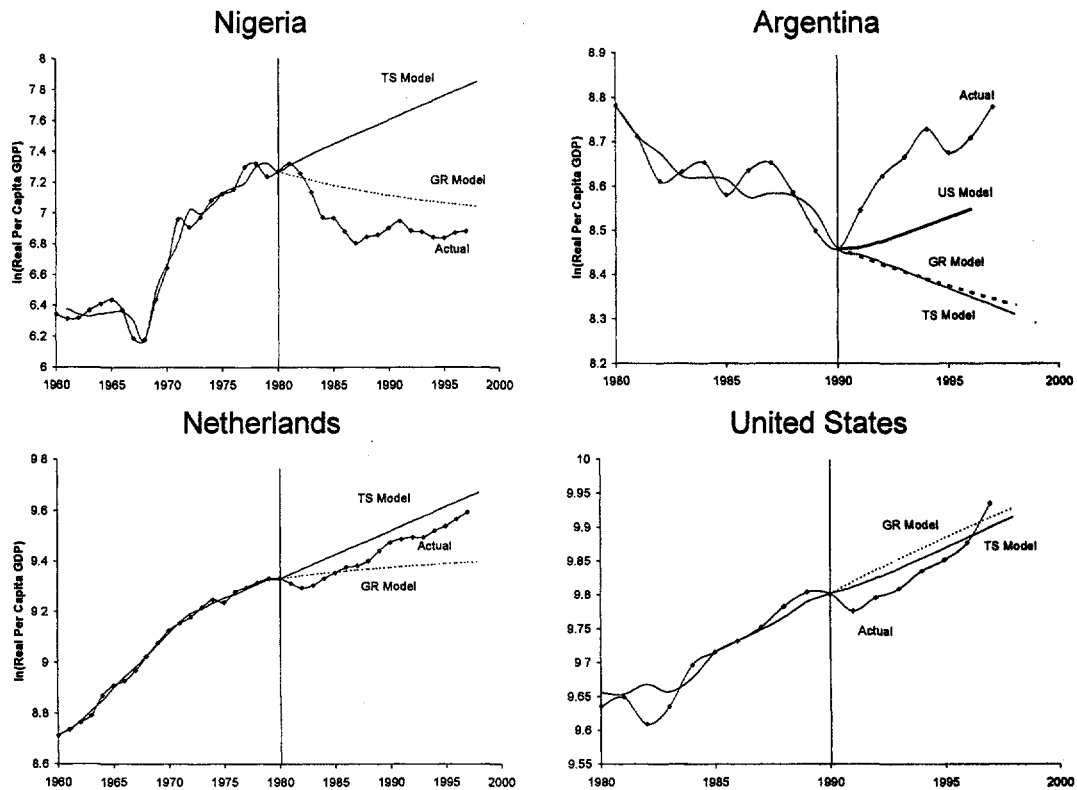
	α	$se(\alpha)$	β	$se(\beta)$	P-Value for:	
					$H_0: \beta=0$	$H_0: \beta=1$
Forecast Origin = 1980						
1981	-0.015	0.006	0.501	0.179	0.005	0.005
1982	-0.057	0.01	0.527	0.163	0.001	0.004
1983	-0.095	0.016	0.654	0.165	0.000	0.036
1984	-0.112	0.02	0.763	0.154	0.000	0.124
1985	-0.121	0.022	0.71	0.138	0.000	0.036
1986	-0.124	0.026	0.783	0.132	0.000	0.100
1987	-0.132	0.029	0.83	0.127	0.000	0.181
1988	-0.133	0.032	0.799	0.122	0.000	0.099
1989	-0.136	0.036	0.804	0.121	0.000	0.105
1990	-0.14	0.039	0.821	0.117	0.000	0.126
1991	-0.15	0.04	0.838	0.109	0.000	0.137
1992	-0.164	0.044	0.854	0.108	0.000	0.176
1993	-0.177	0.047	0.88	0.106	0.000	0.258
1994	-0.177	0.05	0.897	0.105	0.000	0.327
1995	-0.166	0.052	0.917	0.101	0.000	0.411
1996	-0.149	0.055	0.934	0.099	0.000	0.505
1997	-0.164	0.057	0.89	0.099	0.000	0.267
Forecast Origin = 1990						
1991	-0.004	0.005	0.603	0.274	0.028	0.147
1992	-0.013	0.009	0.563	0.237	0.018	0.065
1993	-0.016	0.012	0.797	0.221	0.000	0.358
1994	-0.007	0.013	0.766	0.181	0.000	0.196
1995	0.004	0.014	0.807	0.157	0.000	0.219
1996	0.021	0.017	0.842	0.152	0.000	0.299
1997	0.028	0.017	0.885	0.135	0.000	0.394

Notes: This table reports the results from estimating the following regression:

$$y_{i,t+s} = \alpha + \beta \cdot \hat{y}_{i,t+s|t}^{GR} + (1-\beta) \cdot \hat{y}_{i,t+s|t}^{TS} + u_{i,t+s}$$

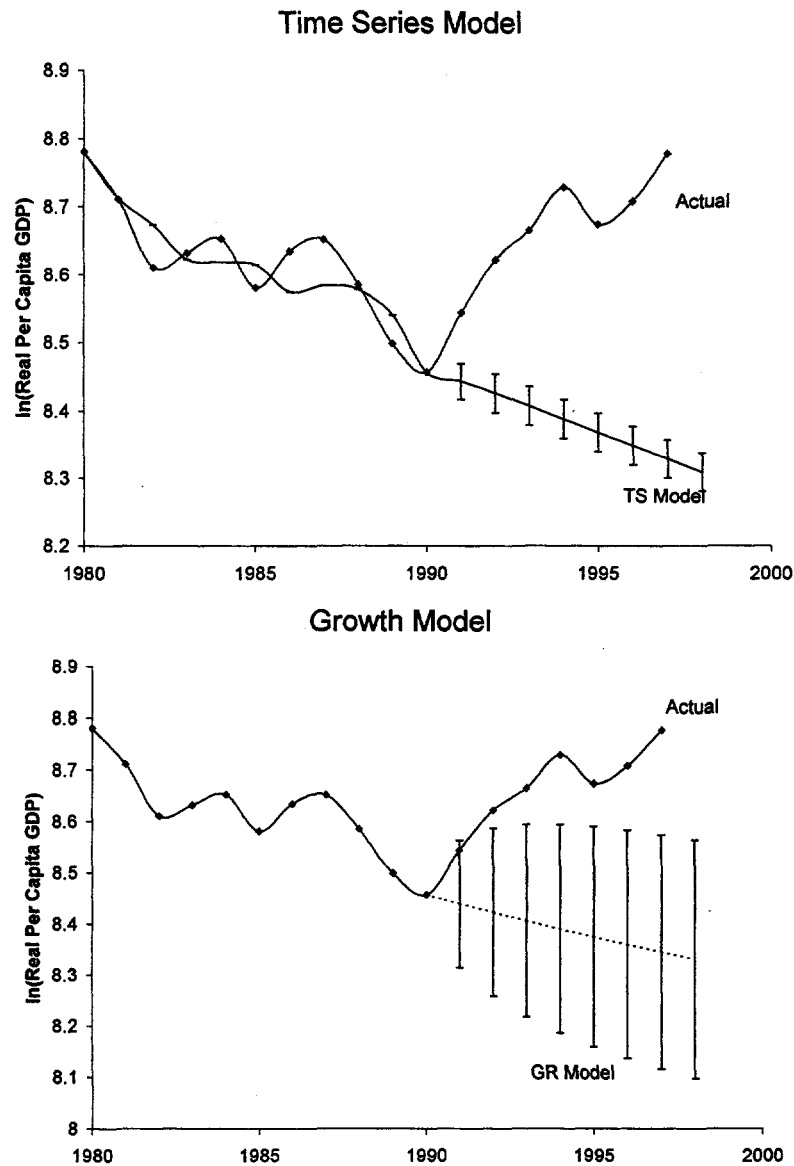
cross-sectionally for each of the indicated years. The last two columns report the p-values corresponding to the null hypothesis that the time series model encompasses the growth model ($\beta=0$) and that the growth model encompasses the time series model ($\beta=1$).

**Figure 1:
A Look at Individual Country Forecasts**



Notes: This figure plots predicted and actual real per capita GDP for the indicated countries. The vertical line in each graph indicates the end of the sample period over which the model was estimated.

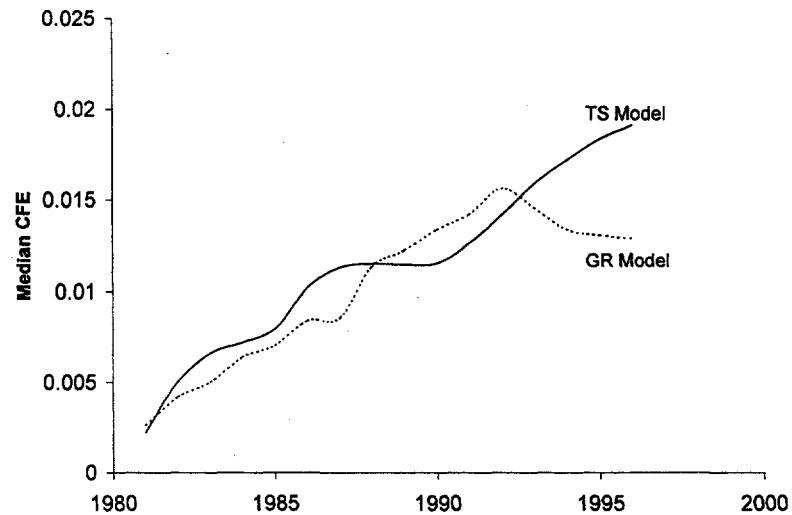
Figure 2:
Forecast Confidence Intervals for Argentina



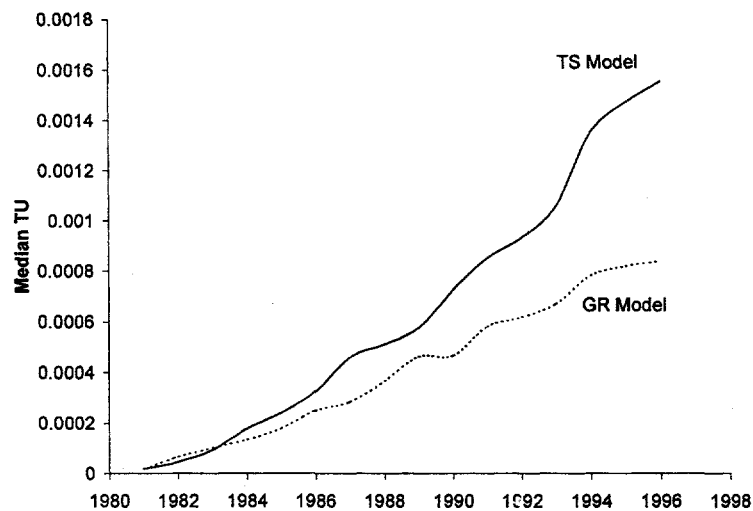
Notes: This figure plots actual and predicted real per capita GDP for Argentina, using information available through 1990. The vertical bars indicate a 90% forecast confidence interval.

Figure 3:
Evaluating Forecast Performance
(Forecasts based on information available through 1980)

Median Cumulative Forecast Errors



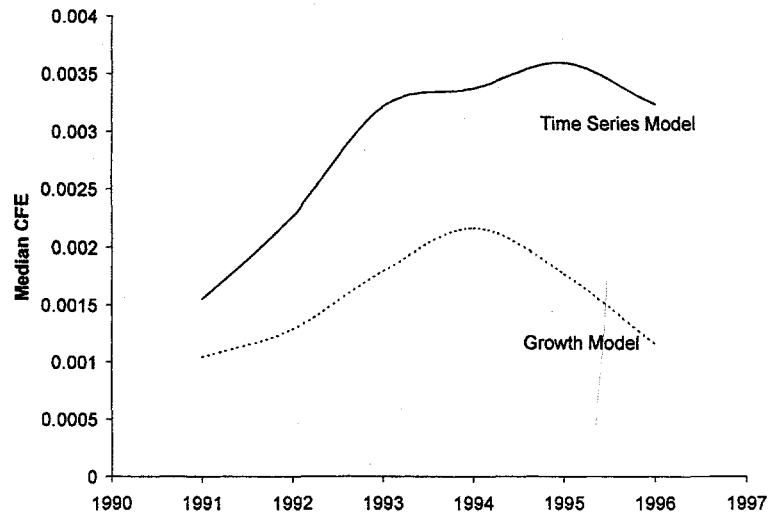
Median Theil-U Statistics



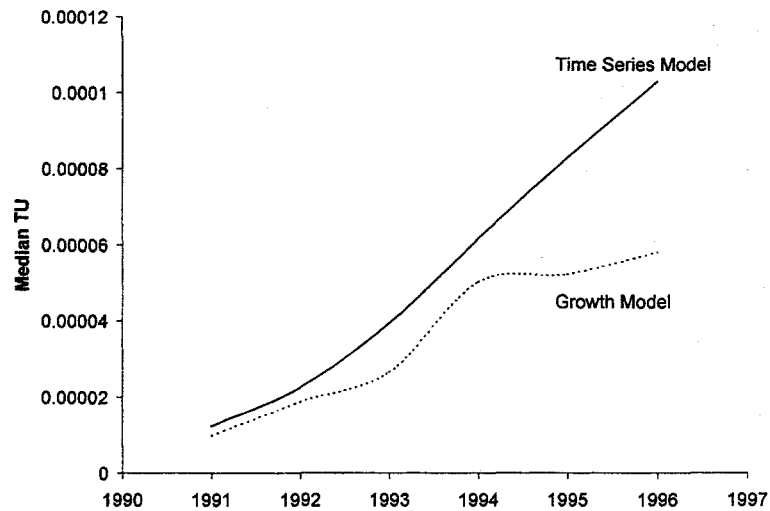
Notes: These graphs report the median cumulate forecast error and Theil U-statistic at each forecast horizon. Medians are taken across the set of 73 countries for which both forecasts are available.

Figure 4:
Evaluating Forecast Performance
(Forecasts based on information available through 1990)

Median Cumulative Forecast Errors



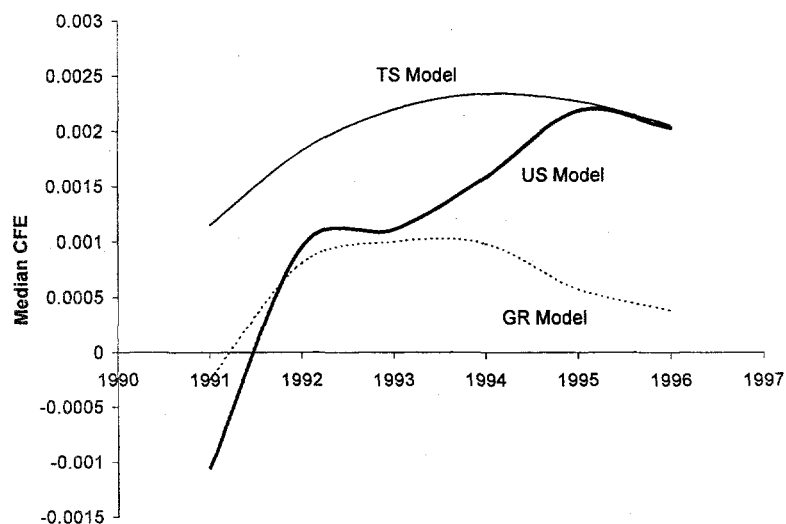
Median Theil-U Statistics



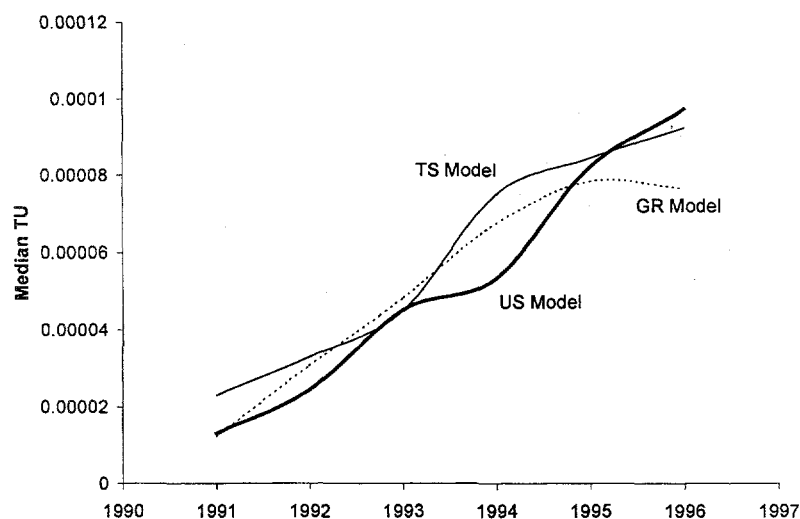
Notes: These graphs report the median cumulate forecast error and Theil U-statistic at each forecast horizon. Medians are taken across the set of 82 countries for which both forecasts are available.

Figure 5:
Evaluating Forecast Performance
(Forecasts based on information available through 1990,
Sample of developing countries for which Unified Survey forecasts are available)

Median Cumulative Forecast Errors



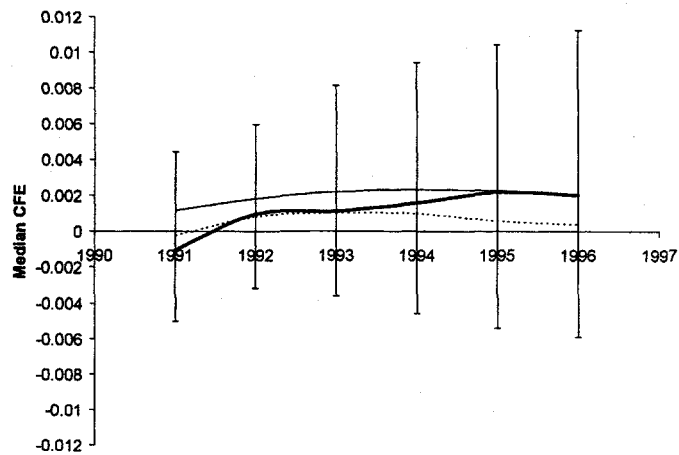
Median Theil-U Statistics



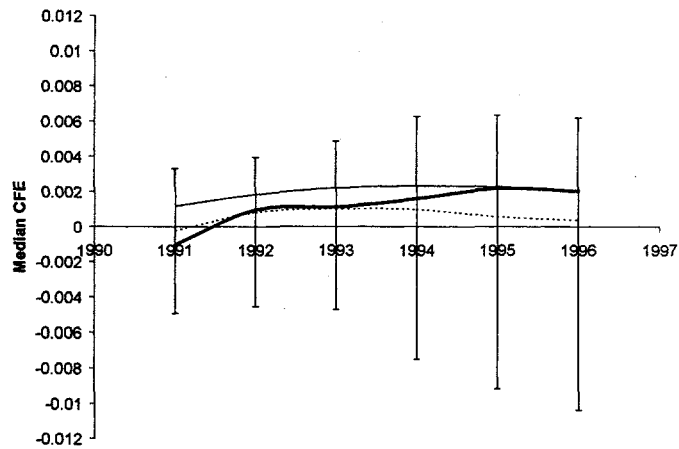
Notes: These graphs report the median cumulate forecast error and Theil U-statistic at each forecast horizon. Medians are taken across the set of 53 countries for which all three forecasts are available.

Figure 6:
Are Differences in Forecast Performance Significant?
(Interquartile Range of Cumulative Forecast Errors)

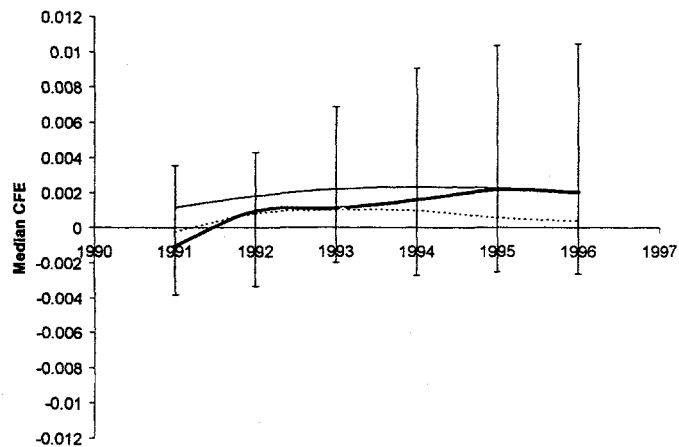
Time Series Model



Growth Model



Unified Survey



Notes: These graphs indicate the median, first, and third quartiles of the cumulative forecast errors associated with the three models, for the set of 53 countries for which these forecasts are available.

Policy Research Working Paper Series

	Title	Author	Date	Contact for paper
WPS2201	Financial Development and Industrial Capital Accumulation	Biagio Bossone	October 1999	E. Mekhova 85984
WPS2202	Specialization without Regret: Transfer Rights, Agricultural Productivity, and Investment in an Industrializing Economy	Michael R. Carter Yang Yao	October 1999	M. Fernandez 33766
WPS2203	Market versus Administrative Reallocation of Agricultural Land in a Period of Rapid Industrialization	Michael R. Carter Yang Yao	October 1999	M. Fernandez 33766
WPS2204	Corruption under Moral Hazard	Gunnar S. Eskeland Henrik Thiele	October 1999	H. Sladovich 37698
WPS2205	Foreign-Owned Capital and Endogenous Tariffs	Marcelo Olarreaga	October 1999	L. Tabada 36896
WPS2206	Household Childcare Choices and Women's Work Behavior in Russia	Michael M. Lokshin	October 1999	P. Sader 33902
WPS2207	Jamaica's Food Stamp Program: Impacts on Poverty and Welfare	Kene Ezemenari Kalanidhi Subbarao	October 1999	G. Peralta 37405
WPS2208	Ethnic Partition as a Solution to Ethnic War: An Empirical Critique of the Theoretical Literature	Nicholas Sambanis	October 1999	H. Sladovich 37698
WPS2209	Does Corruption Relieve Foreign Investors of the Burden of Taxes and Capital Controls?	Shang-Jin Wei	October 1999	H. Sladovich 37698
WPS2210	The Slippery Slope: Explaining the Increase in Extreme Poverty in Urban Brazil, 1976–96	Francisco H. G. Ferreira Ricardo Paes de Barros	October 1999	G. Ilogon 33732
WPS2211	Competition Policy, Developing Countries, and the World Trade Organization	Bernard Hoekman Peter Holmes	October 1999	L. Tabada 36896
WPS2212	Is African Manufacturing Skill-Constrained?	Howard Pack Christina Paxson	October 1999	H. Sladovich 37698
WPS2213	Fiscal Solvency and Sustainability in Economic Management	Hinh T. Dinh	October 1999	M. Rani 32057

Policy Research Working Paper Series

Title	Author	Date	Contact for paper
WPS2214 Trade Policy and Market Access Issues for Developing Countries: Implications for the Millennium Round	Constantine Michalopoulos	October 1999	L. Tabada 36896
WPS2215 Implementation of Uruguay Round Commitments: The Development Challenge	J. Michael Finger Philip Schuler	October 1999	L. Tabada 36896
WPS2216 Corruption and Trade Tariffs, or a Case for Uniform Tariffs	Roberta Gatti	October 1999	R. Gatti 38735
WPS2217 Border, Border, Wide and Far, How We Wonder What You Are	David C. Parsley Shang-Jin Wei	November 1999	H. Sladovich 37698
WPS2218 Who Avoids and Who Escapes Poverty during the Transition: Evidence from Polish Panel Data, 1993-96	Wlodzimierz Okrasa	November 1999	S. Fallon 38009
WPS2219 The Effect of the United States' Granting Most Favored Nation Status to Vietnam	Emiko Fukase Will Martin	November 1999	L. Tabada 36896
WPS2220 A Quantitative Evaluation of Vietnam's Accession to the ASEAN Free Trade Area	Emiko Fukase Will Martin	November 1999	L. Tabada 36896
WPS2221 The Dynamics of Poverty and the Effectiveness of Poland's Safety Net (1993-96)	Wlodzimierz Okrasa	November 1999	S. Fallon 38009
WPS2222 Labor Market Integration in the Presence of Social Capital	Maurice Schiff	November 1999	L. Tabada 36896
WPS2223 Integrated Financial Supervision: Lessons from Northern European Experience	Michael Taylor Alex Fleming	November 1999	S. Torres 39012

11/24